Ensemble Representation of Uncertainty in Lagrangian Satellite Rainfall Estimates

T. J. BELLERBY

Department of Geography, University of Hull, Hull, United Kingdom

(Manuscript received 9 August 2012, in final form 22 April 2013)

ABSTRACT

A new algorithm called Lagrangian Simulation (LSIM) has been developed that enables the interpolation uncertainty present in Lagrangian satellite rainfall algorithms such as the Climate Prediction Center (CPC) morphing technique (CMORPH) to be characterized using an ensemble product. The new algorithm generates ensemble sequences of rainfall fields conditioned on multiplatform multisensor microwave satellite data, demonstrating a conditional simulation approach that overcomes the problem of discontinuous uncertainty fields inherent in this type of product. Each ensemble member is consistent with the information present in the satellite data, while variation between members is indicative of uncertainty in the rainfall retrievals. LSIM is based on the combination of a Markov weather generator, conditioned on both previous and subsequent microwave measurements, and a global optimization procedure that uses simulated annealing to constrain the generated rainfall fields to display appropriate spatial structures. The new algorithm has been validated over a region of the continental United States and has been shown to provide reliable estimates of both point uncertainty distributions and wider spatiotemporal structures.

1. Introduction

Current high-resolution satellite precipitation products combine information from multiple satellite sensors carried on a range of different satellite platforms (Adler et al. 2000; Behrangi et al. 2010; Bellerby et al. 2009; Huffman et al. 2007; Joyce et al. 2004; Kidd et al. 2003; Marzano et al. 2004; Nicholson et al. 2003a,b; Sorooshian et al. 2000; Tapiador et al. 2004; Todd et al. 2001; Turk and Miller 2005; Ushio et al. 2009; Xu et al. 1999). By integrating diverse datasets, these techniques aim to exploit the strengths and minimize the limitations of individual sensors, satellite platforms, and their associated orbits. However, sampling limitations, retrieval uncertainties, and heterogeneities in input data combine to create complex error characteristics in the resulting precipitation estimates. These uncertainties are correlated in both space and time (Bellerby and Sun 2005; Hossain and Anagnostou 2006; Teo 2006; Zeweldi and Gebremichael 2009). The significance of such correlations is demonstrated by Nijssen and Lettenmaier (2004), who employed Monte Carlo simulations of spatially correlated and uncorrelated satellite rainfall errors to estimate the impact of spatial and temporal sampling on hydrological responses, concluding that spatial correlation had a noticeable impact on the relationship between error and basin size, although this effect was relatively less for streamflow and evapotranspiration than for precipitation.

Numerous studies have attempted to quantify the magnitudes of various components of the uncertainty present in satellite rainfall estimates. However, the representation of rainfall uncertainty in satellite products has received less attention. Perhaps the most obvious way of representing this uncertainty is through an additive and/or multiplicative error model:

$$R_{\text{obs}}(\mathbf{x}, t) = \varepsilon_m(\mathbf{x}, t) R_{\text{est}}[S(\mathbf{x}, t)] + \varepsilon_a(\mathbf{x}, t), \qquad (1)$$

where $R_{\text{obs}}(\mathbf{x}, t)$ is the true (observed) rainfall field at location $\mathbf{x}$ and time t; R_{est} is the estimated rainfall based on satellite inputs S; and ε_a and ε_m are additive and multiplicative errors, respectively. However, this formulation is relatively ineffective at handling the intermittent nature of the rainfall field where many errors rest on rain/no-rain misidentification. In addition, it is quite possible for different satellite inputs to produce the same rain-rate estimates but to be associated with very different error structures, causing both ε_a and ε_m to

Corresponding author address: T. J. Bellerby, Department of Geography, University of Hull, Cottingham Road, Hull HU6 7RX, United Kingdom.
E-mail: t.j.bellerby@hull.ac.uk

DOI: 10.1175/JHM-D-12-0121.1

become highly dependent on the input data S. An alternative approach is to consider the conditional distribution of observed rainfall with respect to the satellite inputs (Bellerby and Sun 2005; Bellerby 2007; Gebremichael and Krajewski 2005; Gebremichael et al. 2011):

$$P_u(R;\mathbf{x},t) = P[R_{\text{obs}}(\mathbf{x},t) \mid S(\mathbf{x},t)]. \quad (2)$$

This uncertainty representation incorporates both rainfall intermittency and dependencies on satellite inputs in a straightforward manner. It is also independent of the precise retrieval algorithm employed; it is simply characterized by the input data used. Of course, any practical implementation of Eq. (2) will have to model $P_u(R)$, and this modeling process may incorporate its own assumptions. While useful for many applications, point conditional distributions do not characterize spatiotemporal correlations in rainfall uncertainty. However, these spatial and temporal dependencies may be incorporated into an uncertainty model by using the point conditional distributions as the basis of an ensemble approach.

Ensemble approaches use stochastic weather generators to derive multiple precipitation fields from the same satellite inputs. Each field represents an equiprobable realization of the time series of the rainfall fields, consistent with the available input data while containing a random element commensurate with the associated uncertainty. Geostatistical models are used to ensure that each ensemble member displays a realistic spatiotemporal structure. These representations effectively incorporate spatiotemporal correlations in rainfall uncertainty and integrate readily with deterministic hydrological models (Bowler et al. 2006; Hossain et al. 2004; Hossain and Anagnostou 2006; Nijssen and Lettenmaier 2004; Nikolopoulos et al. 2010.). Ensemble rainfall products have been developed for a range of technologies, including rain gauge networks (Clark and Slater 2006) and terrestrial radar (McMillan et al. 2011). AghaKouchak et al. (2009) generated ensemble simulations conditioned on the Tropical Rainfall Measuring Mission (TRMM) Multi-Satellite Precipitation Analysis (TMPA) product using spatiotemporally uncorrelated additive and multiplicative and additive error fields. Bellerby and Sun (2005) and Teo and Grimes (2007) independently developed conditional ensemble products using local conditional distributions, as per Eq. (2), derived for single-sensor satellite data. These techniques incorporated spatial and temporal dependencies in uncertainty but assumed that these dependencies are spatially and temporally continuous. This assumption does not hold for general multisensor, multiplatform satellite products where the uncertainty field, $P_u(R; \mathbf{x}, t)$, is typically discontinuous in both space and time. There is therefore a need to develop a more generalized approach capable of dealing with current satellite algorithms and datasets.

Most current satellite precipitation products are primarily dependent on passive and active microwave (MW) data. These sensors are sensitive to precipitation-related hydrometeors but are restricted to low Earth orbiting (LEO) platforms necessarily associated with low temporal sampling. Most operational LEO platforms visit a given location twice a day, although data from multiple platforms may be aggregated to yield higher effective sampling frequencies. The forthcoming Global Precipitation Measurement (GPM) mission plans to coordinate a cluster of LEO satellites to provide a maximum 3-h return time across the globe (Hou et al. 2008). Given the sparse and intermittent sampling provided by LEO MW satellite sensors, there has been considerable effort invested in combining these data with other inputs, predominantly infrared (IR) imagery from geostationary Earth orbiting (GEO) satellites, to generate products with higher temporal resolutions and reduced sampling errors. GEO IR data are available every 30-min or better. However, they only show the presence of clouds and are not directly sensitive to precipitation processes. Lagrangian or ''morphing'' algorithms use GEO or other data to estimate the precipitation advection field and then move, and possibly modify, rainfall patterns between LEO satellite overpasses, effectively interpolating rainfall between successive MW measurements along advection streamlines. The longest standing Lagrangian satellite rainfall estimation technique is the operational Climate Prediction Center (CPC) morphing technique (CMORPH; Joyce et al. 2004). CMORPH operates by linearly interpolating rainfall rates between MW overpasses along advection streamlines derived from cloud-tracking GEO IR imagery. Since rainfall advection does not always move in lockstep with cloud motion, an empirical adjustment is employed to derive the former from the latter. Recent variations of CMORPH employ other types of advection information, including model winds (Joyce et al. 2010), and further developments of the concept modify precipitation rates along the streamlines through reference to GEO IR brightness temperatures (Joyce and Xie 2011; Kubota et al. 2007; Ushio et al. 2009) or classified cloud types [Rain Estimation Using Forward-Adjusted Advection of Microwave Estimates (REFAME); Behrangi et al. 2010]. The Lagrangian Model algorithm (LMODEL; Bellerby et al. 2009; Hsu et al. 2009) runs a simplified storm mass balance model along advection streamlines, forced using GEO cloud

development information and corrected against MW overpass data.

A range of studies have considered the implications of LEO sampling on averaged or accumulated products (Astin 1997; Bell and Kundu 1996; Chang and Chiu 2001; Chelton and Schlax 1991; Hong et al. 2006; Roca et al. 2010; Steiner et al. 2003; Zeng and Levy 1995). Uncertainty in the CMORPH algorithm has been studied at a range of spatial and temporal scales and compared to other operational approaches (Anagnostou et al. 2010; Dinku et al. 2008; Ebert et al. 2007; Sapiano and Arkin 2009; Tian et al. 2007; Zeweldi and Gebremichael 2009). Gebremichael et al. (2011) derive conditional distributions of observed rainfall with respect to 3-h, 0.25° CMORPH estimates using a combination of discrete rain/no rain probabilities and a gamma distribution for positive rainfall. This approach proved effective in quantifying a homogenous relationship between satellite estimates and observations. However, the authors note that CMORPH rainfall retrieval uncertainties are dependent on factors other than the precipitation estimate itself, including local satellite sampling. Zeweldi and Gebremichael (2009) found that errors in 1–6-h CMORPH products averaged over a small ($\sim$602 km^2) watershed demonstrated significant temporal correlations.

Given the presence of varying sampling uncertainties and significant spatiotemporal correlations, there is a clear case for developing an ensemble approach to represent the uncertainty in high-resolution Lagrangian satellite rainfall algorithms. It is important to note that such an algorithm should quantify retrieval uncertainty within the satellite product itself using only data available to the algorithm. This differs from investigative studies of satellite rainfall uncertainty that employ high-resolution data, such as terrestrial radar to generate rainfall ensembles (e.g., Hossain et al. 2004; Hossain and Anagnostou 2006; Nikolopoulos et al. 2010). Lagrangian ensemble rainfall simulations conditioned on MW data present some significant difficulties, primarily as a result of discontinuous uncertainty fields. As MW overpasses are morphed along advection streamlines, they combine to form a complex mosaic. Different regions within this mosaic will be computed with reference to different MW overpasses associated with different temporal lags and significantly different error characteristics.

This paper describes the development of a new conditional precipitation simulation algorithm, the Lagrangian Simulation (LSIM), designed to quantify the interpolation uncertainty in an MW-morphing satellite rainfall product. A stochastic rainfall generator is run along GEO-derived advection streamlines conditioned on both previous and subsequent MW measurements. Spatial structures are maintained by imposing stationary covariance functions on the rainfall fields generated at each time step. The following sections detail the structure of the new algorithm and report its validation using MW and surface radar data over a region of the continental United States.

2. Methodology

a. Dataset

The LSIM algorithm was developed using two datasets obtained from the CPC: the 4-km composite global GEO IR dataset and the combined 8-km microwave dataset (MWCOMB). The former is a 30-min composite of available geostationary TIR ($\sim$11 μm) imagery, with individual satellite contributions corrected for zenith angle dependence to reduce interplatform discontinuities (Janowiak et al. 2001). MWCOMB is a composite dataset incorporating data from the Defense Meteorological Satellite Program (DMSP) Special Sensor Microwave Imager (SSM/I), the Polar Operational Environmental Satellite (POES) Advanced Microwave Sounding Unit B (AMSU-B), the *Aqua* Advanced Scanning Microwave Radiometer for Earth Observing System (AMSR-E), and the Tropical Rainfall Measuring Mission (TRMM) Microwave Imager (TMI) instruments. The MWCOMB product is interpolated to a common 8-km spatial resolution and 30-min temporal resolution (Ferraro 1997; Ferraro et al. 2000; Kummerow et al. 2001; Weng et al. 2003). This spatial resolution is finer than that provided by some of the sensors used to generate the composite. To avoid interpolation artifacts complicating the analysis, the MWCOMB data were aggregated to a 0.24° (approximately 24 km) spatial resolution. Data were extracted for a window covering 80.09°–110°W, 20.07°–55°N, the largest box show in Fig. 1. All dataset times were processed in UTC.

An additional dataset was constructed from 0.01°, 2.5-min Q2 radar rainfall data obtained from the National Oceanic and Atmospheric Administration (NOAA) National Severe Storms Laboratory (NSSL; Amitai et al. 2012; Lakshmanan et al. 2007; Vasiloff et al. 2007). These data were aggregated to a 24-km, 30-min spatiotemporal resolution consistent with the aggregated MW data.

b. Algorithm overview

The LSIM algorithm is designed to provide a conditional simulation of the interpolation uncertainty present in a Lagrangian satellite rainfall estimate. Lagrangian techniques temporally interpolate rainfall along estimated advection streamlines. Thus, if microwave sensor estimates R_{mw} are available at times t_1 and t_2, then the

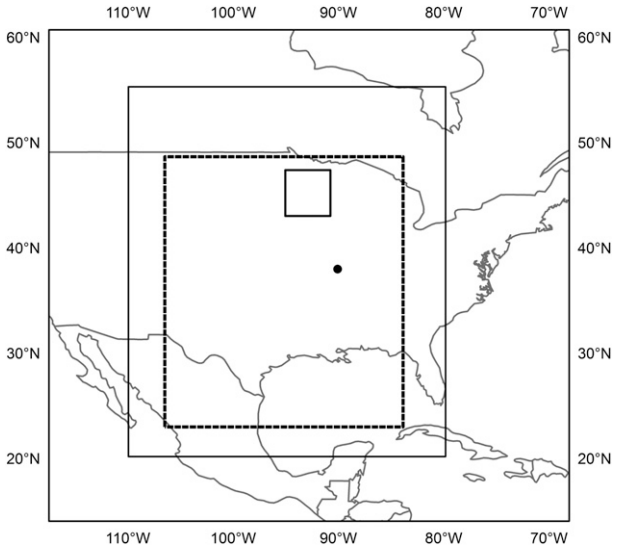

FIG. 1. Location of the study area. The largest box shows the extent of the dataset used to drive the algorithm. The smaller dashed box shows the extent of the data displayed in Fig. 2, and the smallest box shows the window shown in Fig. 4. The point shows the location associated with the plots in Fig. 7.

estimated rainfall $R_{\text{morp}}(\mathbf{x}, t)$ for location $\mathbf{x}$ at an intermediate time t, $t_1 < t < t_2$, is given by

$$R_{\text{morp}}(\mathbf{x}, t) = f[R_{\text{mw}}(\mathbf{x}_{t\to t_1}, t_1), R_{\text{mw}}(\mathbf{x}_{t\to t_2}, t_2), t - t_1, t_2 - t], \quad (3)$$

where f is the interpolation function used by the algorithm and $\mathbf{x}_{t\to t'}$ is the location that would be reached by moving from location $\mathbf{x}$ at time t along the estimated advection field to a corresponding location at previous or subsequent time t'. It should be noted that these algorithms necessarily operate a few hours behind real time since they cannot compute the rainfall for any given point until a subsequent microwave observation is available. Real-time implementations extrapolate forward from the final microwave overpass but update estimates once they are bracketed by both a previous and subsequent observation.

For the general Lagrangian algorithm described by Eq. (3), the conditional distribution $P_u(R; \mathbf{x}, t)$ of observed rainfall with respect to the satellite information used to derive the estimate is given by

$$P_u(R; \mathbf{x}, t) = P[R_{\text{obs}} \mid R_{\text{mw}}(\mathbf{x}_{t\to t_1}, t_1), R_{\text{mw}}(\mathbf{x}_{t\to t_2}, t_2), t - t_1, t_2 - t]. \quad (4)$$

Here $P(R \mid \ldots)$ is a conditional distribution of rainfall characterized only by the previous and subsequent microwave rain-rate observations and their respective lags $t - t_1$ and $t_2 - t$. It is important to clarify the meaning of "observed rainfall" when discussing this conditional distribution. In LSIM, observations are equated with microwave satellite measurements, meaning that the algorithm is assessed with respect to its ability to interpolate between satellite overpasses to derive intermediate values as if they had been measured by a microwave sensor. Inaccuracies in the individual microwave measurements themselves are thus not modeled. This enables LSIM products to be both calibrated and generated using satellite data alone. Microwave retrieval errors and their possible incorporation into LSIM are further discussed in section 4.

To derive an ensemble precipitation product, it is necessary to randomly sample points from the conditional distributions in Eq. (4) while generating rainfall fields that have the correct spatial and temporal structures. Temporal structure is imposed by building up each ensemble member time step by time step, with each field directly conditioned on its predecessor. Spatial structure is imposed by imposing geostatistical constraints on the fields generated at each time step. This is implemented by minimizing a cost function using an optimization procedure based on simulated annealing.

c. Determining precipitation advection

The current LSIM implementation derives precipitation advection from high-resolution (4 km) GEO IR imagery using the multiresolution deformed-mesh tracking algorithm of Bellerby (2006). This algorithm associates each cloudy pixel in an IR image with a corresponding pixel in a previous image to a demonstrated accuracy of ~2–3 pixels. The algorithm forms the basis of a number of published Lagrangian satellite precipitation algorithms (Behrangi et al. 2010; Bellerby et al. 2009). The algorithm was applied at full GEO image resolution and then subsampled to the 24-km product resolution.

d. Determining temporal dependencies

Conditional rainfall probabilities are initially determined for discrete rainfall categories and then interpolated to yield a continuous conditional distribution function (cdf). LSIM defines $N + 1$ rainfall categories $R_k = \{R \mid L_k \le R < L_{k+1}\}$, $k = 0 \ldots N$, with respective lower bounds L_k; R_0 is the zero-rainfall category; and R_N has no upper bound. This formulation enables a transition between rainfall states over a single time step to be modeled using a stationary $N \times N$ transition matrix:

$$T_{kj} = P[R(\mathbf{x}, t) \in R_k \mid R(\mathbf{x}_{t\to t-\delta t}, t - \delta t) \in R_j], \quad (5)$$

where δt is the time step. Working along the streamline, probabilities at each time step must be conditioned on both the simulation for the previous time step and the subsequent microwave observation at time t_2. This gives rise to the following equation for category conditional probabilities (Sherlaw-Johnson et al. 1995):

$$P[R \in R_k \mid R(\mathbf{x}_{t\to t-\delta t}, t-\delta t) \in R_i, R_{\mathrm{mw}}(\mathbf{x}_{t\to t_2}, t_2) \in R_j] = \frac{T_{k,i}(T^m)_{j,k}}{(T^{m+1})_{i,j}}. \quad (6)$$

Once the category probabilities have been derived for a given location and time, the complete conditional probability function, covering the full continuous range of rainfall rates, is estimated using

$$P_u(R;\mathbf{x},t) = P(R \mid R \in R_k) \times P[R \in R_k \mid R(\mathbf{x}_{t\to t-\delta t}, t-\delta t), R_{\mathrm{mw}}(\mathbf{x}_{t\to t_2}, t_2)]. \quad (7)$$

The Markov transition matrix $T_{j,k}$ must be obtained from the available satellite data. These data are intermittent, with a varying number of time steps between satellite overpasses. Transition matrices may be obtained from data with nonuniform time intervals using an expectation maximization (EM) approach (Sherlaw-Johnson et al. 1995). This EM algorithm proceeds as follows:

Step 1—Initialization: The transition matrix $T_{j,k}$ is given an initial value. This may be derived from MW data separated by a single time step, although in the absence of any such data $T_{j,k}$ may be assigned an arbitrary nonzero value.

Step 2—The E step: Assume that two MW measurements are m time steps apart and lie within rain-rate categories R_i and R_j. The probability $P_{k,l}$ that a rainfall category transition from R_k to R_l will occur between time step n and time step $n + 1$ within this interval is given by

$$P_{l,k} = \frac{(T^n)_{k,i} T_{l,k} (T^{m-n-1})_{j,l}}{(T^n)_{j,i}}. \quad (8)$$

This equation may be used to interpolate multistep data to single time steps to build up a cross histogram of single-step transitions. Note that counts in this histogram will not necessarily be whole numbers since Eq. (8) assigns transition probabilities rather than identifying actual transitions.

TABLE 1. Rain-rate categories used by the LSIM algorithm.

R_0	$R = 0\,\mathrm{mm\,h^{-1}}$
R_1	$0\,\mathrm{mm\,h^{-1}} < R \le 0.5\,\mathrm{mm\,h^{-1}}$
R_2	$0.5\,\mathrm{mm\,h^{-1}} < R \le 1.0\,\mathrm{mm\,h^{-1}}$
R_3	$1.0\,\mathrm{mm\,h^{-1}} < R \le 2.0\,\mathrm{mm\,h^{-1}}$
R_4	$2.0\,\mathrm{mm\,h^{-1}} < R \le 5.0\,\mathrm{mm\,h^{-1}}$
R_5	$5.0\,\mathrm{mm\,h^{-1}} < R \le 10.0\,\mathrm{mm\,h^{-1}}$
R_6	$10.0\,\mathrm{mm\,h^{-1}} < R \le 20.0\,\mathrm{mm\,h^{-1}}$
R_7	$R > 20.0\,\mathrm{mm\,h^{-1}}$

Step 3—The M step: Once the transition histogram has been determined from all available measurements, it is used to compute a new transition matrix $T_{j,k}$. If this matrix differs significantly from the former value of $T_{j,k}$ (with a maximum absolute difference greater than 10^{-8} for any element) the algorithm repeats from Step 2. Otherwise, it terminates.

Note that the EM algorithm is applied along advection streamlines, which means that the transition matrix $T_{j,k}$ depends on the choice of advection scheme. The strength of the temporal dependence represented by $T_{j,k}$ could be used to compare different schemes; a more effective scheme will be associated with a stronger temporal relationship.

Table 1 shows the rainfall categories employed by the prototype implementation of LSIM. These categories were chosen as a compromise between covering the range of rain rates to a sufficient resolution and ensuring that all categories had sufficient membership to enable the transition matrix to be determined. These categories are used to derive full continuous probability distributions via Eq. (7). Table 2 shows the derived transition matrix $T_{j,k}$, which is assumed to apply across the study area and period. Note that some cross probabilities are as low as 10^{-4}. It is thus essential to use a large dataset to derive $T_{j,k}$.

e. Spatial structure

At each time step the LSIM algorithm must generate a random spatial field that simultaneously satisfies the point conditional distributions given by Eq. (7) and displays the required spatial structure. This is achieved by minimizing a cost function using simulated annealing. Simulated annealing is a global optimization technique that may be used as the basis of effective conditional simulation algorithms in a range of contexts (Goovaerts 1997), including precipitation simulation (Bardossy 1998; Haberlandt and Gattke 2004). The LSIM implementation is based on the approach of Dafflon et al. (2009). An initial rainfall field is constructed by independently drawing each pixel rainfall rate from its

TABLE 2. Rainfall category transition matrix for one 30-min time step used by the Markov model.

		Rainfall category at previous time step							
		R_0	R_1	R_2	R_3	R_4	R_5	R_6	R_7
Rainfall category at current time step	R_0	0.9778	0.0141	0.0035	0.0024	0.0016	0.0004	0.0001	0.0001
	R_1	0.6181	0.1636	0.0671	0.0660	0.0577	0.0195	0.0058	0.0021
	R_2	0.4335	0.1893	0.0932	0.1128	0.1088	0.0441	0.0129	0.0055
	R_3	0.3183	0.1871	0.0985	0.1436	0.1562	0.0653	0.0251	0.0058
	R_4	0.1832	0.1385	0.0887	0.1466	0.2451	0.1223	0.0590	0.0168
	R_5	0.0890	0.0908	0.0662	0.1057	0.2547	0.1973	0.1497	0.0465
	R_6	0.0367	0.0335	0.0382	0.0687	0.2416	0.2365	0.2096	0.1353
	R_7	0.0122	0.0180	0.0122	0.0376	0.1461	0.2097	0.2447	0.3194

local conditional distribution given by Eq. (7). Pixels are then selected at random and a potential replacement rainfall value is randomly selected from the conditional distribution for that pixel. The replacement is accepted if it reduces an objective function O, which quantifies the deviation from the required spatial structure. If the new value would increase O, the replacement is randomly accepted or rejected, with an acceptance probability P_a given by

$$P_a = e^{[-(O_{\mathrm{NEW}} - O_{\mathrm{OLD}})]/T}, \tag{9}$$

where T is the annealing temperature. In the early stages of the algorithm T is given a relatively large value associated with a high probability for accepting positive changes to O (determined by examining a sample of possible replacements). As the algorithm progresses, T is steadily reduced. Thus, in the early stages of the algorithm, there is a significant chance of the state ''jumping'' out of local minima, while in later stages the algorithm is allowed to converge to a globally optimal state. Convergence is determined by a failure to improve the objective by more than a set convergence criterion in a given number of time steps (respectively set in the prototype implementation to 0.001 and 10 times the number of pixels in the window being simulated).

The objective function contains four elements, respectively scaled by multipliers $\mu_1 \ldots \mu_4$:

$$O = \mu_1 O_1 + \mu_2 O_2 + \mu_3 O_3 + \mu_4 O_4. \tag{10}$$

The first two elements quantify the difference between the simulated and observed (MW) covariance and rain/no-rain indicator covariance functions, respectively:

$$O_1 = \sum_h |C^{\mathrm{SIM}}(h) - C^{\mathrm{MW}}(h)|$$

$$O_2 = \sum_h |C_I^{\mathrm{SIM}}(h) - C_I^{\mathrm{MW}}(h)|,$$

$$C(h) = \mathrm{Cov}[R(\mathbf{x}), R(\mathbf{x} + \delta\mathbf{x}) \,|\, \|\delta\mathbf{x}\| = h],$$

$$C_I(h) = \mathrm{Cov}\{I[R(\mathbf{x})], I[R(\mathbf{x} + \delta\mathbf{x})] \,|\, \|\delta\mathbf{x}\| = h\},$$

$$I(R) = \begin{cases} 0 & R = 0 \\ 1 & R > 0 \end{cases}. \tag{11}$$

Observed variograms, calculated for 24-km increments in distance, are employed directly rather than being modeled, facilitating a fully automatic and rapid calibration procedure. To avoid the long-range effects present in empirical variograms, both $C(h)$ and $C_I(h)$ are truncated when their values fall below 1% of that at $h = 0$. The scaling terms μ_1 and μ_2 are calculated to give $\mu_1 O_1$ and $\mu_2 O_2$ unit initial values. The use of empirical variograms enables LSIM calibration to be fully automatic.

The combination of rainfall-rate and indicator covariance functions is generally effective at controlling the wider structure of the rainfall field, but it does permit some admixture of point, or very small area, low-rainfall-rate noise. This effect is countered by introducing a penalty term based on the number of raining pixels $N(\mathbf{x})$ immediately adjacent to the pixel under consideration:

$$O_3 = \sum_{\mathbf{x}} \mathrm{Max}\left(\frac{1}{P\{I[R(\mathbf{x})] \,|\, N(\mathbf{x})\}} - 2, 0\right), \tag{12}$$

where empirical conditional probabilities $P[I(R) \,|\, N]$ have been derived from the MW dataset. The penalty term for an individual pixel is zero if its rain/no-rain state is more likely than the alternative and increases toward infinity as the conditional probability of the indicator function $I(R)$ tends toward zero. This penalty term can interact problematically with O_2 in the initial stages of the simulated annealing process, preventing the necessary introduction of intermediate rain/no-rain structures. To prevent this, μ_3 is set to zero until the optimization has converged. It is then scaled to give $\mu_3 O_3$ a small positive value (0.1), and the optimization is allowed to rapidly reconverge.

The final term of the objective function is designed to maintain adherence to the conditional probabilities of Eq. (7). The simulated annealing process already incorporates the conditional distributions when drawing new random pixel values. However, the geostatistical constraints may cause selective sampling, causing some drift away from the expected category proportions. The corresponding objective term is computed by noting that if for a given rainfall category R_k, M pixels are associated with the conditional probability P, then $N = MP$ of these pixels should have simulated rainfall rates in R_k. To implement this as a cost function, conditional probabilities for each category k are divided into discrete bins $p = 1, 2\ldots$ associated with respective mean conditional probabilities $\overline{P}_p$, giving

$$O_4 = \sum_k \sum_p |N_{p,k} - \overline{P}_p M_{p,k}|. \tag{13}$$

To cater for conditional probabilities that are small with respect to the number of pixels in the image, a random rounding scheme is applied to $\overline{P}_p M_{p,k}$. This scheme will, for example, round 5.2 to 5.0 for 80% of the time and to 6.0 for 20% of the time. The initial realization of the simulated field will closely adhere to the required conditional probabilities; μ_4 is therefore set to scale the initial value of $\mu_4 O_4$ to one-half of the objective function convergence criterion. This permits convergence providing the value of O_4 does not become significantly worse than its initial value.

3. Results

The LSIM algorithm was used to generate a 100-member precipitation ensemble for July 2011. A simple ''linear morphing'' product was also generated for the same period. This comparison product was generated by linearly interpolating rainfall rates between MW satellite overpasses along the same IR-derived streamlines. The comparison product will be referred to LMORPH in the discussion below. It should be noted that while this product is structurally similar to CMORPH, it differs from the operational algorithm in a number of respects, including advection estimation. To provide a satellite-based validation dataset, MW data for the specific times of day, as follows, were excluded from the input data used to generate both the LSIM and LMORPH products: 0115, 0315, 0745, 1045, 1445, 1745, 1915, and 2045 UTC (the remaining 30-min intervals in each day were used as input to the LSIM algorithm). These times were selected to provide a broad spread of temporal lags between validation overpasses and both preceding and subsequent MW measurements. While an exact time of day tends to be associated with the same satellite overpass, this is not invariably the case and day to day variations serve to further vary the sampling.

As an example of the LSIM product, Fig. 2 shows the development of a part of the rainfall field between two MW overpasses spanning 1845–2015 UTC on 3 July 2011, as modeled by three arbitrarily selected ensemble members. Each ensemble member is designed to be commensurate with the input satellite data while variation between ensembles is indicative of the remaining uncertainty. The broad structure of the rainfall field remains the same in each member, while smaller-scale details vary. This is commensurate with studies that note that the LEO sampling characterizes larger rainfall structures more effectively than smaller ones (Bellerby 2012), resulting from the longer persistence of larger structures (Germann and Zawadzki 2002, 2004). This effect is important. While the LSIM covariance model is capable of describing precipitating structures at a range of spatial scales, the generated precipitation fields are conditioned only on local MW overpasses and this includes the size and complexity of simulated storm systems. In the absence of conditioning data, storm size and complexity would be randomly sampled from the distribution observed over the complete study area and period. Moreover, an isotropic model would not preferentially generate linear structures such as frontal systems. However, a manual review of LSIM outputs suggests that larger-scale structures are effectively constrained by the available MW data throughout test dataset.

An ensemble product is reliable if the fractional occurrence of a given event across the ensemble matches its frequency of occurrence in coincident observed data (Toth et al. 2003). Within an M-member ensemble, it is possible to identify all times and locations at which exactly N members simulate rainfall in a given category. If the product is absolutely reliable, then an exact proportion M/N of these times and locations will be associated with observed rainfall in that category. Figure 3 compares simulated and coincident observed exceedence probabilities for seven rainfall thresholds corresponding to LSIM category boundaries and for the zero-rainfall category. As simulated exceedence probabilities increase (or as zero-rainfall probabilities decrease), the chance of that probability occurring in the ensemble product declines dramatically. To avoid low sampling rates obscuring the comparison, a minimum cutoff of 500 observations was applied when generating the reliability plots. This process excluded different proportions of points depending on the rainfall rate threshold applied. Approximately 5% of observations were excluded from the $>0\,\mathrm{mm\,h^{-1}}$ comparison, ranging down

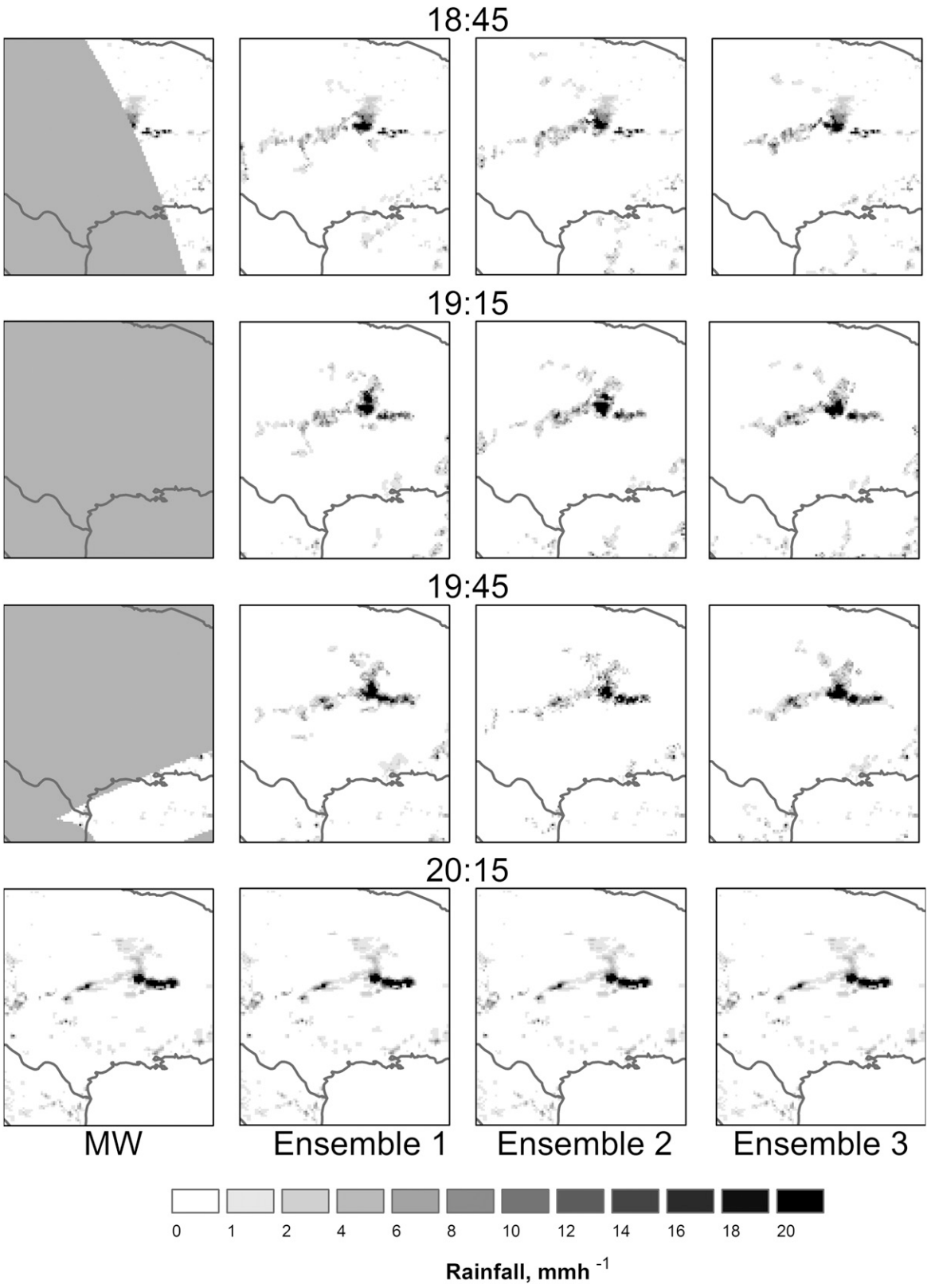

FIG. 2. Example ensemble product showing the development of a single storm between microwave satellite overpasses on 3 July 2011 within the region bounded by the smaller dashed box in Fig. 1. Ensemble members were arbitrarily selected. All times are UTC.

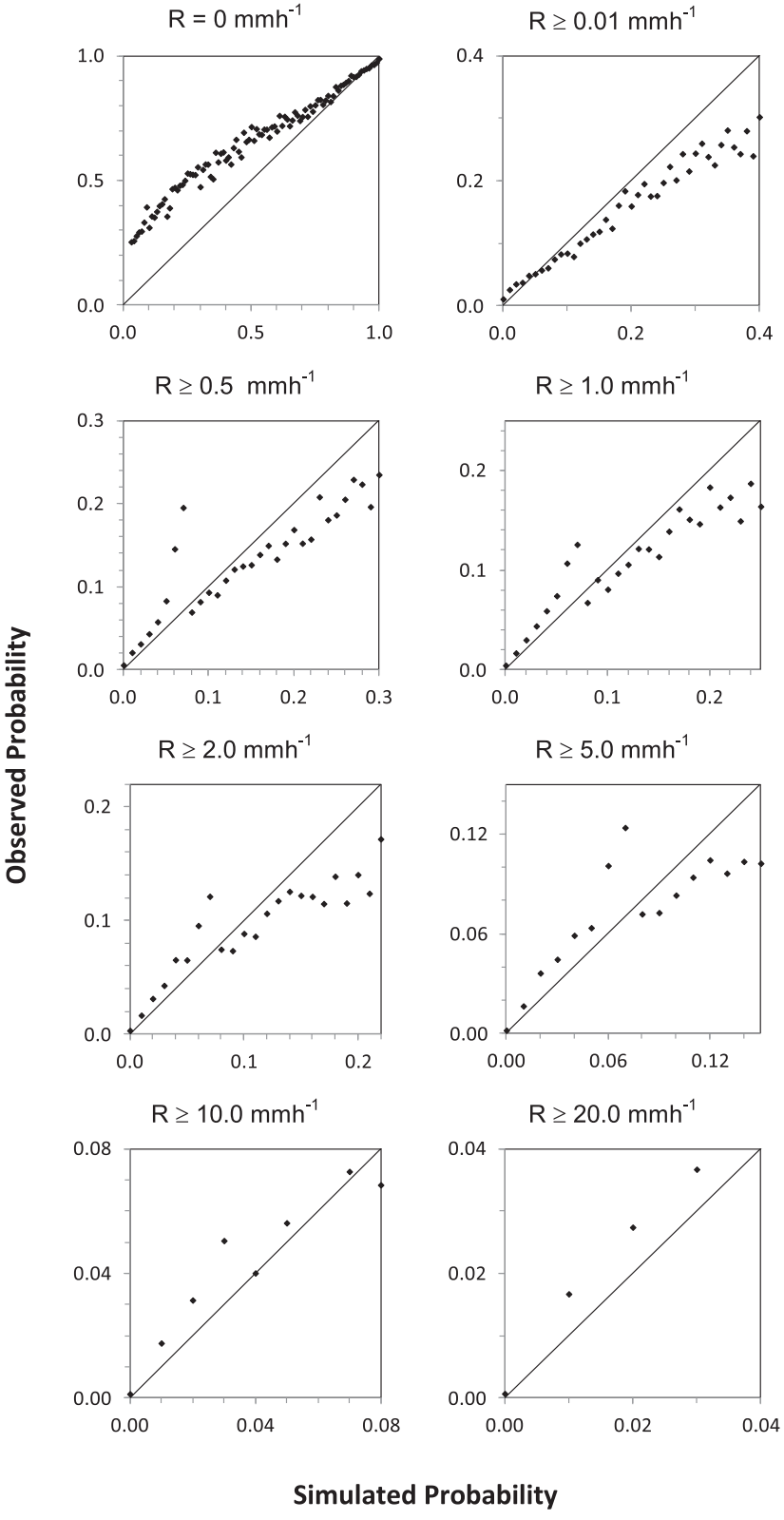

FIG. 3. Reliability plots of the 24-km ensemble product compared against microwave satellite data from the validation dataset. Each plot compares exceedence probabilities determined from the LSIM ensemble to coincident observed probabilities in the validation dataset.

to 0.9% of observations excluded from the $>20\,\mathrm{mm\,h^{-1}}$ plot. The cutoff was not applied to the zero-rainfall category plot, enabling the low-number sampling effects to be seen. There is generally a good match between simulated and observed probabilities, although the rain/no-rain boundary proves the most difficult to characterize.

Table 3 displays validation statistics based on a comparison of observed and simulated rainfall category probabilities. Note that, as with Fig. 3, categories are defined using single thresholds rather than rainfall rate ranges, as used by the LSIM probability model. Statistics are provided for a range of ensemble sizes and also for point probabilities produced through a direct application of Eq. (7) without reference to the ensemble field generation process. Correlations are a slightly unstable metric to employ in this context, but they do demonstrate the strength of the predicative relationship using a familiar measure. Root-mean-square errors (RMSEs) provide a more robust measure of the probability match. These are clearly dependent on ensemble size, reducing to ~3%–4% for a 100-member ensemble. The 100-member RMSEs are not significantly worse than those provided by the underlying probability model, suggesting that the choice of ensemble size is not inappropriate, particularly given the high computation cost of generating very large ensembles. Biases are generally of the order of 0.1%–0.4%.

Figure 4 compares a small area of the rainfall field for 1515 UTC 15 July 2011 for two ensemble members, the linear morphing product and the input MW satellite data. There are some clear differences between the LSIM and LMORPH outputs, particularly with respect to rainfall rates in the feature to the top left of the window. It is also notable that the morphing product displays a linear discontinuity corresponding to the edge of the MW swath. Morphing products are not spatially incoherent; their skill at rainfall estimation gives their outputs some reasonable spatial structure. However, discontinuities in their underlying uncertainty structure can translate into spatial and temporal discontinuities in the products themselves. By contrast, LSIM rainfall fields are constrained to display consistent spatial structures.

Figure 5 compares a range of quantitative structural attributes for both LSIM and LMORPH products to the MW satellite data. Figure 5a plots the temporal correlation of rainfall rates along advection streamlines as a function of time difference. There is a marked contrast between the correlation curves for the LMORPH and LSIM. Linear interpolation introduces significantly too much autocorrelation along advection streamlines. The LSIM outputs provide a closer match to the Lagrangian

TABLE 3. Validation statistics for LSIM-estimated rainfall category probabilities compared to coincident observational frequencies independent MW satellite data. Statistics are given for a range of ensemble sizes and for point conditional probabilities calculated using the underlying probability model.

	Correlation				RMSE				Bias			
	10 member ensemble	50 member ensemble	100 member ensemble	Point conditional probabilities	10 member ensemble	50 member ensemble	100 member ensemble	Point conditional probabilities	10 member ensemble	50 member ensemble	100 member ensemble	Point conditional probabilities
$R > 0\,\mathrm{mm\,h^{-1}}$	0.993	0.994	0.992	0.976	0.059	0.047	0.047	0.032	0.003	0.003	0.003	−0.018
$R > 0.5\,\mathrm{mm\,h^{-1}}$	0.995	0.995	0.994	0.988	0.050	0.040	0.040	0.021	0.004	0.004	0.004	−0.013
$R > 1\,\mathrm{mm\,h^{-1}}$	0.997	0.996	0.994	0.990	0.047	0.038	0.038	0.017	0.004	0.004	0.004	−0.01
$R > 2\,\mathrm{mm\,h^{-1}}$	0.999	0.994	0.990	0.994	0.043	0.035	0.036	0.011	0.003	0.003	0.003	−0.006
$R > 5\,\mathrm{mm\,h^{-1}}$	0.997	0.977	0.973	0.993	0.037	0.032	0.032	0.006	0.002	0.002	0.002	−0.003
$R > 10\,\mathrm{mm\,h^{-1}}$	0.977	0.943	0.916	0.960	0.033	0.030	0.030	0.007	0.001	0.001	0.001	−0.002
$R > 20\,\mathrm{mm\,h^{-1}}$	0.944	0.847	0.794	0.887	0.024	0.022	0.022	0.005	0.001	0.001	0.001	−0.001

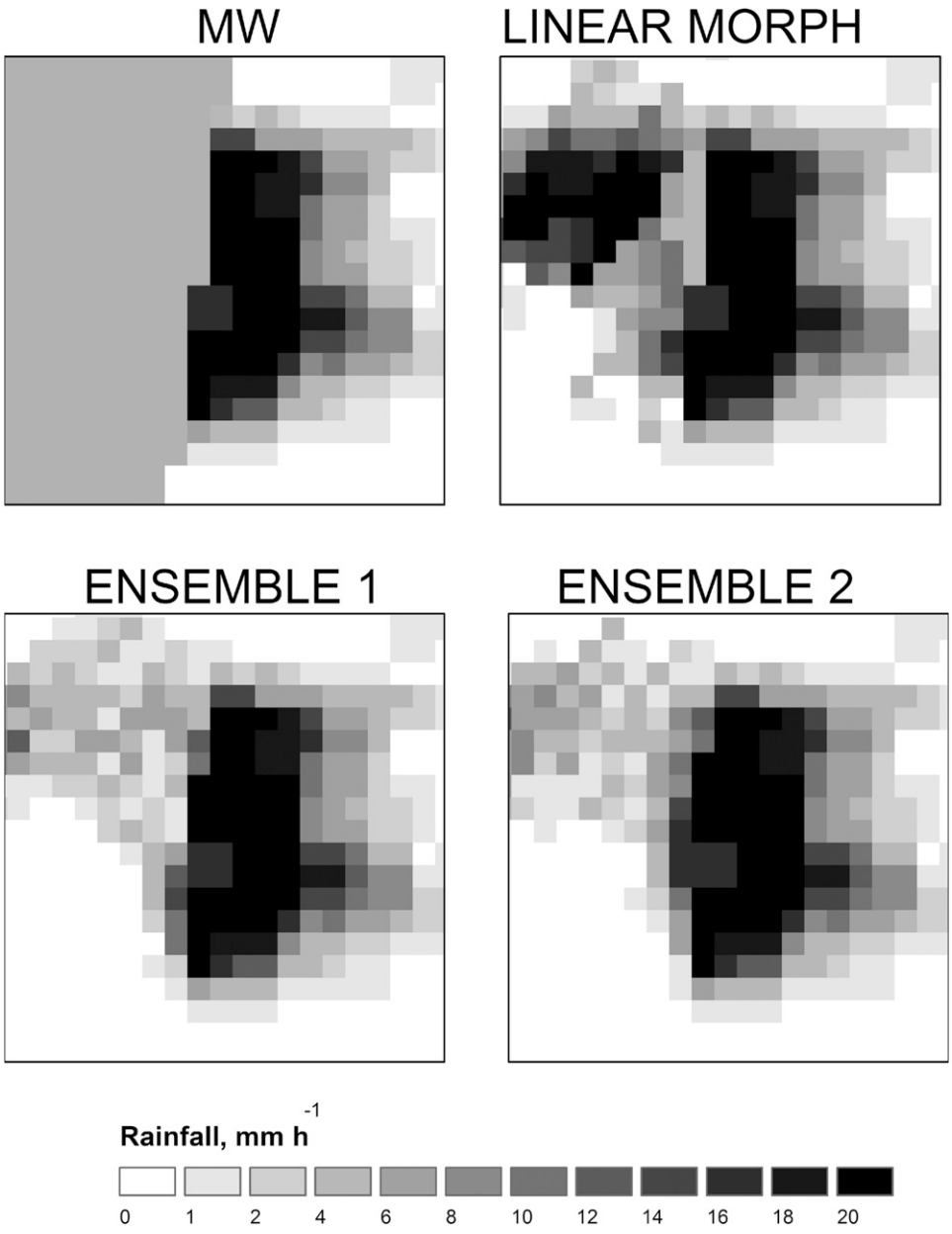

FIG. 4. Example rainfall fields for 1515 UTC 15 Jul 2011 for two arbitrary members of the LSIM ensemble, together with coincident MW satellite data and a corresponding linear morphing product. The area mapped corresponds to the smallest box in Fig. 1.

temporal structure of the MW data. The remaining discrepancies are attributable to the differences between a single-step Markov model of temporal structure and a multistep (and thus multiscale) correlation model. Figure 5b plots temporal correlation in a fixed, Eulerian reference frame. Here the contrast between the LSIM and LMORPH outputs is less significant, and there is a much closer match between the LSIM and MW correlation curves.

The performance of LSIM ensemble members as rainfall estimates was assessed though a validation against 30-min Q2 surface radar data. Table 4 shows validation statistics for LSIM and LMORPH products, computed at a range of spatial and temporal resolutions. LSIM outputs display consistently lower correlations and higher RMSE compared to the linear morphing product. Each member of the LSIM ensemble introduces variability into the precipitation field in order to generate

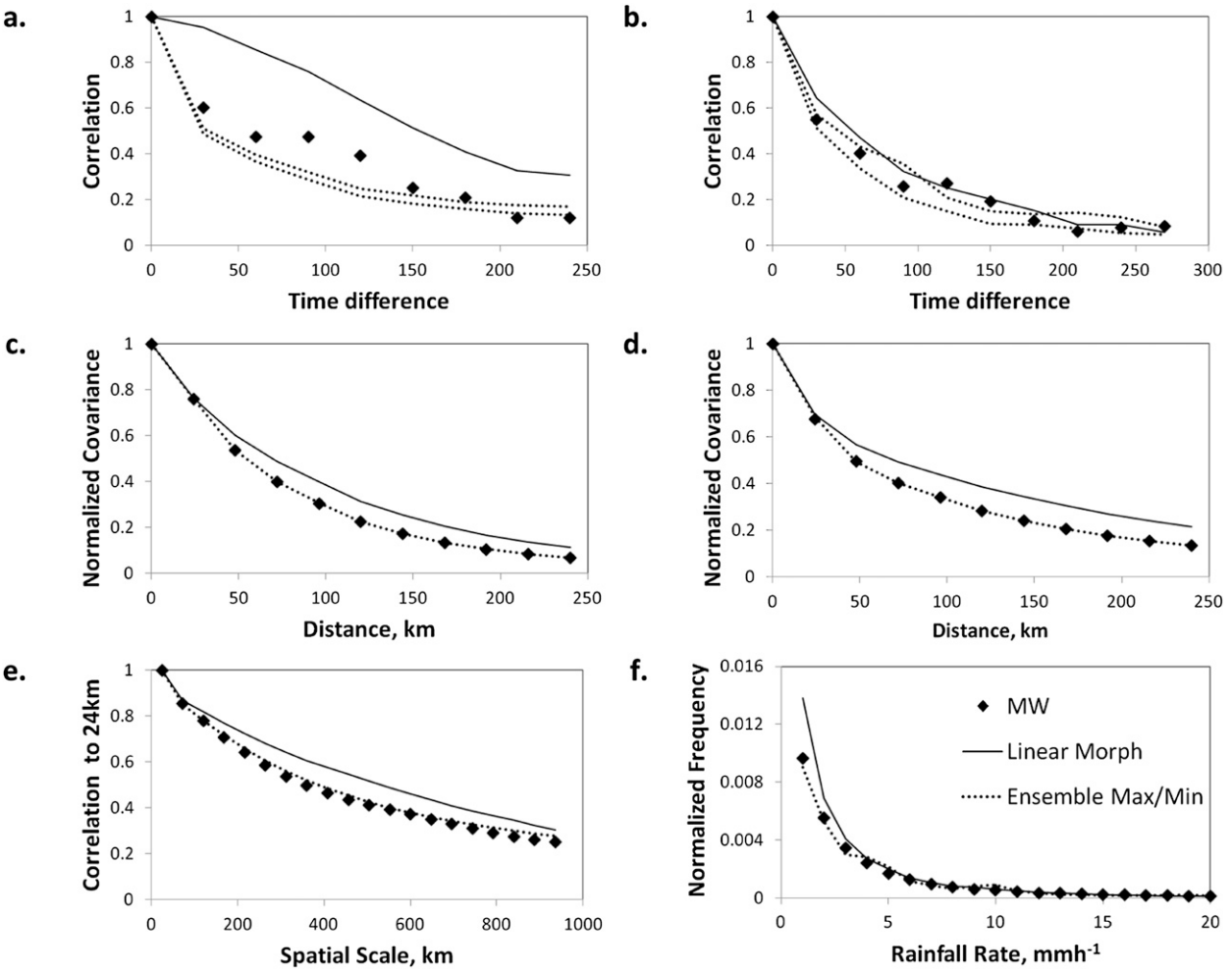

FIG. 5. A comparison of structural characteristics of the LSIM ensemble product to those of the input MW satellite data. Results for a linear morphing product are also shown. (a) Temporal correlation along advection streamlines. (b) Temporal correlation for stationary points. (c) Spatial covariance. (d) Spatial indicator (rain/no rain) covariance. (e) Rainfall-rate correlation across spatial scales. (f) Rainfall rate frequency distribution.

a realistic spatiotemporal structure. Since this variability is stochastically generated, it is unlikely to match the true precipitation field on a point-by-point basis. Statistics are also provided for point-expected values, based on Eq. (7). These may be regarded as the ''best'' rainfall estimate at each point, based on the input data and the assumptions of the probability model. These point estimates outperform the LMORPH product slightly. This suggests that linear interpolation is close to being the optimal interpolation method in this context. However, the very low 30-min rain/no-rain skill score for the expected values is notable. Attaining high point error scores comes at the expense of losing a key structural property: the probability of zero rainfall. Any significant probability of nonzero rainfall will yield a positive expected rainfall value, and few combinations of satellite inputs can guarantee that no rainfall will be present. This emphasizes the observation that satellite rainfall estimation cannot be about point error statistics alone—rainfall structures are also important. Ensemble products offer one way to reconcile these potentially conflicting demands.

Figure 6 plots correlation with surface radar against time to the nearest MW overpass along an advection streamline. The linearly morphed estimates and ensemble expected value match each other closely, with the latter again giving the slightly better performance. Correlations for ensemble members are lower but decay in a similar pattern. These results further emphasize the trade-off between point errors and acceptable precipitation structures. The ensemble expected value optimizes point estimates at the expense of spatial structure,

TABLE 4. Validation statistics against ground radar for LSIM ensemble members, a linear morphing product, and point-expected values computed using the LSIM underlying probability model. Entries for LSIM show the minimum and maximum across the ensemble.

Resolution Number of points	24 km, 30 min 13 225 439			24 km, 180 min 1 171 433			96 km, 180 min 54 613		
Product	Linear morph	Expected values	LSIM ensemble	Linear morph	Expected values	LSIM ensemble	Linear morph	Expected values	LSIM ensemble
Correlation	0.43	0.42	0.24–0.33	0.63	0.65	0.38–0.51	0.76	0.79	0.51–0.65
RMSE ($mm\,h^{-1}$)	1.32	0.99	1.67–1.77	0.81	0.50	1.19–1.28	0.67	0.37	0.99–1.07
Bias ($mm\,h^{-1}$)	0.10	0.08	0.11–0.12	0.11	0.04	0.14–0.15	0.12	0.04	0.14–0.16
Skill (%)	90.3	19.3	89.5–90.9	93.3	94.0	91.5–93.4	93.7	94.8	91.3–93.6

as noted above. Linear morphing closely produces a result close to this point-by-point optimum. Individual ensemble members introduce realistic spatial variation into the spatial field at the expense of a less optimal fit at each point location.

Figure 7 plots the simulated conditional histograms of 3-h, 24-km rainfall totals for a single location throughout 3 July 2011. These histograms were created by dividing ensemble outputs into eight bins, each covering an equal range of rainfall rates. Note that the rainfall rates on the category axis show the maximum value in each bin. Simulated cdf's vary considerably in form from time step to time step. This wide variation in form and the presence of highly skewed uncertainty distributions are observed throughout the dataset. The bins containing the observed radar rainfall rate for each time step are also indicated. These results suggest that the need to move beyond simple error measures such as variance to a more complete representation of retrieval uncertainty in satellite rainfall products persists at the temporal resolutions characteristic of GPM.

4. Conclusions

A new algorithm, LSIM, has been developed to stochastically simulate sequences of rainfall fields conditioned on MW satellite data within a Lagrangian framework defined by GEO-derived advection streamlines. The LSIM algorithm simulates the rainfall field directly rather than replacing the conditional simulation of rainfall with the unconditional simulation of rainfall uncertainty. It is thus able to cope with discontinuities in the uncertainty field, such as those occurring at MW swath boundaries, and is thus more generally applicable than previous methodologies. LSIM simulations have been shown to be reliable, with simulated exceedence probabilities for interpolated rainfall rates matching those encountered in coincident independent MW satellite data, with RMSEs of the order of 3%.

One of the main limitations of the LSIM approach lies in its model of spatiotemporal rainfall structure, which relies on stationary spatial covariance functions and a stationary single-step Markov process. In reality, precipitation will display differing structures at a range of temporal scales, and these structures will vary between meteorological regimes and with forcing factors such as time of day. While the covariance model represents precipitation structures at a range of spatial scales, it does not adopt the structures generated to the prevailing meteorological regime. In the absence of conditioning data, LSIM would generate precipitation structures displaying a range of sizes and complexities according to the broad geostatistics of the calibration region and period. In practice, the satellite data tend to constrain the spatial structures more tightly, as evidenced by the variation between ensemble members in Fig. 2, which is typical of LSIM output as a whole. More complex representations are available, including copula-based

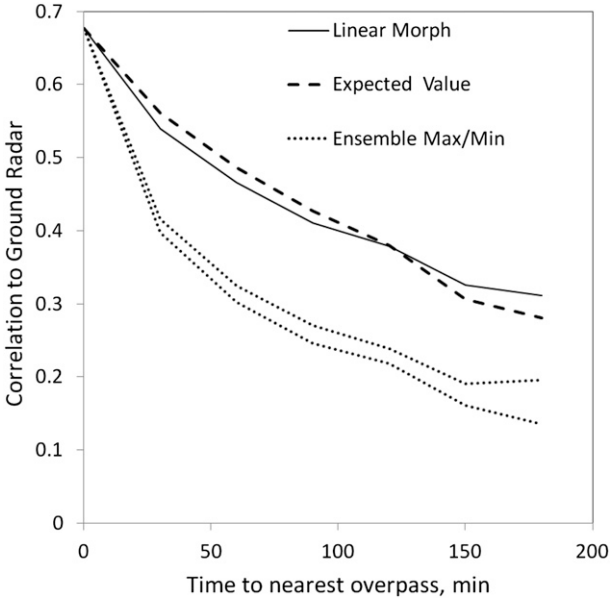

FIG. 6. Comparison of temporal correlation along an advection streamline as a function of time to the nearest microwave satellite overpass for the LSIM ensemble, point-expected values, and a linear morphing product.

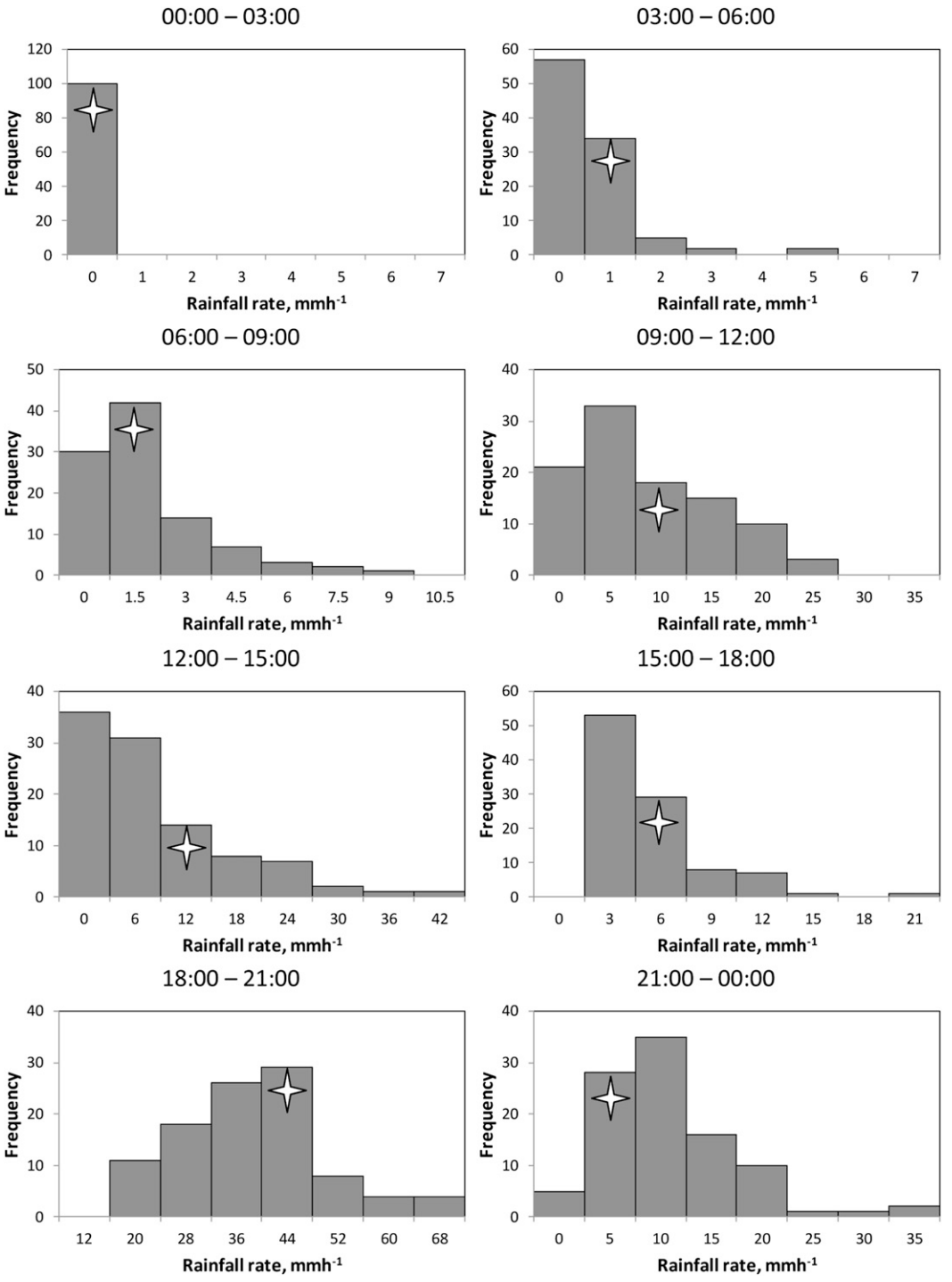

FIG. 7. Conditional distributions of 3-h, 24-km rainfall rates at 90.08°W, 37.96°N (the point shown on Fig. 1) for 3 July 2011 derived from a 100-member LSIM ensemble product. The markers indicate the bin in which the corresponding surface radar rainfall total lies. All times are UTC.

approaches (AghaKouchak et al. 2010). However, if satellite data alone are used to characterize product uncertainty, data availability to calibrate more complex models may be an issue.

LSIM is designed as a satellite-only product and consequently does not incorporate uncertainty in the satellite MW estimates themselves. If a suitably general uncertainty model were available, through physical

modeling or extensive ground validation studies, then it could be incorporated into the algorithm. Definite values for MW samples would be replaced by conditional distributions determined by the retrieval error model through a straightforward extension of Eq. (7). It would also be necessary to ensure that both spatial and temporal models were compatible with the new definition of the ''observed'' rainfall field. However, it should also be noted that MW satellite data are available globally while high-quality ground data are only available in limited areas. An approach that relied on input from ground data may thus lack global applicability.

The LSIM implementation described in this paper uses the same GEO-based advection scheme as the LMODEL and REFAME algorithms. However, other advection schemes (including model reanalysis or LEO-only-based schemes) could be directly substituted into LSIM without modifying its structure. Advection estimation in Lagrangian algorithms is an unresolved issue, with no definitive scheme yet identified (Joyce et al. 2010). Precipitation advection has been shown to be scale dependent (Bellerby 2012), which corresponds to the general finding of Salby (1982a,b) that LEO satellite sampling can only characterize larger Fourier components of the field being monitored. The LSIM approach may be used to evaluate advection schemes; the transition matrix T will become closer to the identity matrix when based on an improved advection field. However, it cannot necessarily be used to overcome fundamental limitations implicit in the point advection approach.

LSIM uses geostationary IR data to derive advection streamlines, but not to modify precipitation rates along those streamlines. To extend LSIM to incorporate GEO cloud development information, it would be necessary to additionally condition simulated rainfall on the geostationary image data. In some cases, this may be achievable using a multivariate contingency table, especially if the GEO input is discrete, such as the cloud categories used in the REFAME algorithm. However, as the number of degrees of freedom increases, so does the data requirement, and functional approaches such as artificial neural networks (Bellerby 2007) may be more appropriate for more complex data combinations.

The LSIM algorithm is computationally demanding, but not impossibly so. The prototype implementation completed a single image for a single ensemble member in approximately 1.5 min on one core of a 3-GHz processor. Global application in real time would thus require the use of several hundred cores, either a small part of a grid engine or a computational workstation employing the 64 or 100 core processor chips now available. Further development of the constrained optimization algorithm would potentially yield significant savings in this area.

Acknowledgments. Data for this study were obtained from the NASA Tropical Rainfall Measuring Mission Science Data and Information System; the Center for Hydrometerology and Remote Sensing, University of California, Irvine, with thanks due to D. Braithwaite; and the NOAA National Severe Storms Laboratory, with thanks due to S. Vasiloff, J. Zhang, and K. Howard.

REFERENCES

Adler, R. F., G. J. Huffman, D. T. Bolvin, S. Curtis, and E. J. Nelkin, 2000: Tropical rainfall distributions determined using TRMM combined with other satellite and gauge information. *J. Appl. Meteor.,* **39,** 2007–2023.

AghaKouchak, A., N. Nasrollahi, and E. Habib, 2009: Accounting for uncertainties of the TRMM satellite estimates. *Remote Sens.,* **1,** 606–619.

——, E. Habibb, and A. Bárdossy, 2010: A comparison of three remotely sensed rainfall ensemble generators. *Atmos. Res.,* **98,** 387–399.

Amitai, E., W. Petersen, X. Llort, and S. Vasiloff, 2012: Multiplatform comparisons of rain intensity for extreme precipitation events. *IEEE Trans. Geosci. Remote Sens.,* **50,** 675–686.

Anagnostou, E. N., V. Maggioni, E. Nikolopoulos, T. Taye, and F. Hossain, 2010: Benchmarking high-resolution global satellite rain products to radar and rain gauge rainfall estimates. *IEEE Trans. Geosci. Remote Sens.,* **48,** 1667–1683.

Astin, I., 1997: A survey of studies into errors in large scale space-time averages of rainfall, cloud cover, sea surface processes and the earth's radiation budget as derived from low Earth orbit satellite instruments because of their incomplete temporal and spatial coverage. *Surv. Geophys.,* **18,** 385–403.

Bardossy, A., 1998: Generating precipitation time series using simulated annealing. *Water Resour. Res.,* **34,** 1737–1744.

Behrangi, A., B. Imam, K. Hsu, S. Sorooshian, T. J. Bellerby, and G. J. Huffman, 2010: REFAME: Rain Estimation Using Forward-Adjusted Advection of Microwave Estimates. *J. Hydrometeor.,* **11,** 1305–1321.

Bell, T. L., and P. K. Kundu, 1996: A study of the sampling error in satellite rainfall estimates using optimal averaging of data and a stochastic model. *J. Climate,* **9,** 1251–1268.

Bellerby, T., 2006: High-resolution 2-D cloud-top advection from geostationary satellite imagery. *IEEE Trans. Geosci. Remote Sens.,* **44,** 3639–3648.

——, 2007: Satellite rainfall satellite rainfall uncertainty estimation using an artificial neural network. *J. Hydrometeor.,* **8,** 1397–1412.

——, 2012: Searchlight: Precipitation advection tracking using multiplatform low-Earth orbiting satellite data. *IEEE Trans. Geosci. Remote Sens.,* **51,** 2177–2187, doi:10.1109/TGRS.2012.2211604.

——, and J. Sun, 2005: Probabilistic and ensemble representations of the uncertainty in an IR/microwave satellite precipitation product. *J. Hydrometeor.,* **6,** 1032–1044.

——, K. Hsu, and S. Sorooshian, 2009: LMODEL: A satellite precipitation methodology using cloud development modeling. Part I: Algorithm construction and calibration. *J. Hydrometeor.,* **10,** 1081–1095.

Bowler, N. E., C. E. Pierce, and A. W. Seed, 2006: STEPS: A probabilistic precipitation forecasting scheme which merges an extrapolation nowcast with downscaled NWP. *Quart. J. Roy. Meteor. Soc.,* **132,** 2127–2155.

Chang, A. T. C., and L. S. Chiu, 2001: Non-systematic errors of monthly oceanic rainfall derived from passive microwave radiometry. *Geophys. Res. Lett.,* **28,** 1223–1226.

Chelton, D. B., and M. G. Schlax, 1991: Estimation of time averages from irregularly spaced observations: With application to coastal zone color scanner estimates of chlorophyll concentration. *J. Geophys. Res.,* **96,** 14 669–14 692.

Clark, M. P., and A. G. Slater, 2006: Probabilistic quantitative precipitation estimation in complex terrain. *J. Hydrometeor.,* **7,** 3–22.

Dafflon, B., J. Irving, and K. Holliger, 2009: Simulated-annealing-based conditional simulation for the local-scale characterization of heterogeneous aquifers. *J. Appl. Geophys.,* **68,** 60–70.

Dinku, T., S. Chidzambwa, P. Ceccato, S. J. Connor, and C. F. Ropelewski, 2008: Validation of high-resolution satellite rainfall products over complex terrain. *Int. J. Remote Sens.,* **29,** 4097–4110.

Ebert, E. E., J. Janowiak, and C. Kidd, 2007: Comparison of near-real-time precipitation estimates from satellite observations and numerical models. *Bull. Amer. Meteor. Soc.,* **88,** 47–64.

Ferraro, R. R., 1997: Special Sensor Microwave Imager derived global rainfall estimates for climatological applications. *J. Geophys. Res.,* **102,** 16 715–16 736.

——, F. Weng, N. C. Grody, and L. Zhao, 2000: Precipitation characteristics over land from the NOAA-15 AMSU sensor. *Geophys. Res. Lett.,* **27,** 2669–2672.

Gebremichael, M., and W. F. Krajewski, 2005: Modeling distribution of temporal sampling errors in area-time-averaged rainfall estimate. *Atmos. Res.,* **73,** 243–259.

——, G.-Y. Liao, and J. Yan, 2011: Nonparametric error model for a high resolution satellite rainfall product. *Water Resour. Res.,* **47,** W07504, doi:10.1029/2010WR009667.

Germann, U., and I. Zawadzki, 2002: Scale-dependence of the predictability of precipitation from continental radar images. Part I: Description of the methodology. *Mon. Wea. Rev.,* **130,** 2859–2873.

——, and ——, 2004: Scale dependence of the predictability of precipitation from continental radar images. Part II: Probability forecasts. *J. Appl. Meteor.,* **43,** 74–89.

Goovaerts, P., 1997: *Geostatistics for Natural Resources Evaluation*. Oxford University Press, 483 pp.

Haberlandt, U., and C. Gattke, 2004: Spatial interpolation v. simulation of precipitation for rainfall-runoff modelling—A case study in the Lippe river basin. *Hydrology: Science and practice for the 21st century,* B. Webb et al., Eds., British Hydrological Society, 120–127.

Hong, Y., K. Hsu, H. Moradkhani, and S. Sorooshian, 2006: Uncertainty quantification of satellite precipitation estimation and Monte Carlo assessment of the error propagation into hydrologic response. *Water Resour. Res.,* **42,** W08421, doi:10.1029/2005WR004398.

Hossain, F., and E. N. Anagnostou, 2006: A two-dimensional satellite rainfall error model. *IEEE Trans. Geosci. Remote Sens.,* **44,** 1511–1521.

——, ——, and T. Dinku, 2004: Sensitivity analyses of satellite rainfall retrieval and sampling error on flood prediction uncertainty. *IEEE Trans. Geosci. Remote Sens.,* **42,** 130–139.

Hou, A. Y., G. Skofronick-Jackson, C. D. Kummerow, and J. M. Shepherd, 2008: Global precipitation measurement. *Precipitation: Advances in Measurement, Estimation, and Prediction,* S. Michaelides, Ed., Springer, 131–170.

Hsu, K., T. Bellerby, and S. Sorooshian, 2009: LMODEL: A satellite precipitation methodology using cloud development modeling. Part II: Validation. *J. Hydrometeor.,* **10,** 1096–1108.

Huffman, G. J., and Coauthors, 2007: The TRMM Multisatellite Precipitation Analysis (TMPA): Quasi-global, multiyear, combined-sensor precipitation estimates at fine scales. *J. Hydrometeor.,* **8,** 38–55.

Janowiak, J. E., R. J. Joyce, and Y. Yarosh, 2001: A real-time global half-hourly pixel-resolution infrared dataset and its applications. *Bull. Amer. Meteor. Soc.,* **82,** 205–217.

Joyce, R. J., and P. Xie, 2011: Kalman filter–based CMORPH. *J. Hydrometeor.,* **12,** 1547–1563.

——, J. E. Janowiak, P. A. Arkin, and P. Xie, 2004: CMORPH: A method that produces global precipitation estimates from passive microwave and infrared data at high spatial and temporal resolution. *J. Hydrometeor.,* **5,** 487–503.

——, P. Xie, Y. Yarosh, J. E. Janowiak, and P. A. Arkin, 2010: CMORPH: A ''morphing''approach for high resolution precipitation product generation. *Satellite Rainfall Applications for Surface Hydrology,* M. Gebremichael and F. Hossain, Eds., 23–37.

Kidd, C., D. R. Kniveton, M. C. Todd, and T. J. Bellerby, 2003: Satellite rainfall estimation using combined passive microwave and infrared algorithms. *J. Hydrometeor.,* **4,** 1088–1104.

Kubota, T., and Coauthors, 2007: Global precipitation map using satelliteborne microwave radiometers by the GSMaP Project: Production and validation. *IEEE Trans. Geosci. Remote Sens.,* **45,** 2259–2275.

Kummerow, C., and Coauthors, 2001: Evolution of the Goddard Profiling Algorithm (GPROF) for rainfall estimation from passive microwave sensors. *J. Appl. Meteor.,* **40,** 1801–1820.

Lakshmanan, V., A. Fritz, T. Smith, K. Hondl, and G. J. Stumpf, 2007: An automated technique to quality control radar reflectivity data. *J. Appl. Meteor.,* **46,** 288–305.

Marzano, F. S., M. Palmacci, D. Cimini, G. Giuliani, and F. J. Turk, 2004: Multivariate statistical integration of satellite infrared and microwave radiometric measurements for rainfall retrieval at the geostationary scale. *IEEE Trans. Geosci. Remote Sens.,* **42,** 1018–1032.

McMillan, H., B. Jackson, M. Clark, D. Kavetski, and R. Woods, 2011: Rainfall uncertainty in hydrological modelling: An evaluation of multiplicative error models. *J. Hydrol.,* **400,** 83–94.

Nicholson, S. E., and Coauthors, 2003a: Validation of TRMM and other rainfall estimates with a high-density gauge dataset for West Africa. Part I: Validation of GPCC rainfall product and pre-TRMM satellite and blended products. *J. Appl. Meteor.,* **42,** 1337–1354.

——, and Coauthors, 2003b: Validation of TRMM and other rainfall estimates with a high-density gauge dataset for West Africa. Part II: Validation of TRMM rainfall products. *J. Appl. Meteor.,* **42,** 1355–1368.

Nijssen, B., and D. P. Lettenmaier, 2004: Effect of precipitation sampling error on simulated hydrological fluxes and states: Anticipating the Global Precipitation Measurement satellites. *J. Geophys. Res.,* **109,** D02103, doi:10.1029/2003JD003497.

Nikolopoulos, E. I., E. N. Anagnostou, F. Hossain, M. Gebremichael, and M. Borga, 2010: Understanding the scale relationships of

uncertainty propagation of satellite rainfall through a distributed hydrologic model. *J. Hydrometeor.,* **11,** 520–532.

Roca, R., P. Chambon, I. Jobard, P.-E. Kirstetter, M. Gosset, and J. C. Bergès, 2010: Comparing satellite and surface rainfall products over West Africa at meteorologically relevant scales during the AMMA campaign using error estimates. *J. Appl. Meteor. Climatol.,* **49,** 715–731.

Salby, M. L., 1982a: Sampling theory for asynoptic satellite observations. Part I: Space–time spectra, resolution, and aliasing. *J. Atmos. Sci.,* **39,** 2577–2600.

——, 1982b: Sampling theory for asynoptic satellite observations. Part II: Fast Fourier synoptic mapping. *J. Atmos. Sci.,* **39,** 2601–2614.

Sapiano, M. R. P., and P. A. Arkin, 2009: An intercomparison and validation of high-resolution satellite precipitation estimates with 3-hourly gauge data. *J. Hydrometeor.,* **10,** 149–166.

Sherlaw-Johnson, C., S. Gallivan, and J. Burridge, 1995: Estimating a Markov transition matrix from observational data. *J. Oper. Res. Soc.,* **46,** 405–410.

Sorooshian, S., K. Hsu, X. Gao, H. V. Gupta, B. Imam, and D. Braithwaite, 2000: Evaluation of PERSIANN system satellite-based estimates of tropical rainfall. *Bull. Amer. Meteor. Soc.,* **81,** 2035–2046.

Steiner, M., T. L. Bell, Y. Zhang, and E. F. Wood, 2003: Comparison of two methods for estimating the sampling-related uncertainty of satellite rainfall averages based on a large radar dataset. *J. Climate,* **16,** 3759–3778.

Tapiador, F. J., C. Kidd, V. Levizzani, and F. S. Marzano, 2004: A neural networks-based fusion technique to estimate half-hourly estimates at 0.1° resolution from satellite passive microwave and infrared data. *J. Appl. Meteor.,* **43,** 576–594.

Teo, C.-K., 2006: Application of satellite-based rainfall estimates to crop yield forecasting in Africa. Ph.D. thesis, University of Reading, 242 pp. [Available online at: http://www.met.rdg.ac.uk/phdtheses/Application%20of%20satellite-based%20rainfall%20estimates%20to%20crop%20yield%20forecasting%20in%20Africa.pdf.]

——, and D. I. F. Grimes, 2007: Stochastic modelling of rainfall from satellite data. *J. Hydrol.,* **346,** 33–50.

Tian, Y., C. D. Peters-Lidard, B. J. Choudhury, and M. Garcia, 2007: Multitemporal analysis of TRMM-based satellite precipitation products for land data assimilation applications. *J. Hydrometeor.,* **8,** 1165–1183.

Todd, M., C. Kidd, D. R. Kniveton, and T. J. Bellerby, 2001: A combined satellite infrared and passive microwave technique for the estimation of small scale rainfall. *J. Atmos. Oceanic Technol.,* **18,** 742–755.

Toth, Z., O. Talagrand, G. Candille, and Y. Zhu, 2003: Probability and ensemble forecasts. *Forecast Verification,* I. T. Jolliffe and D. B. Stephenson, Eds., Wiley, 137–164.

Turk, F. J., and S. D. Miller, 2005: Toward improving estimates of remotely-sensed precipitation with MODIS/AMSR-E blended data techniques. *IEEE Trans. Geosci. Remote Sens.,* **43,** 1059–1069.

Ushio, T., and Coauthors, 2009: A Kalman filter approach to the Global Satellite Mapping of Precipitation (GSMaP) from combined passive microwave and infrared radiometric data. *J. Meteor. Soc. Japan,* **87A,** 137–151.

Vasiloff, S. V., and Coauthors, 2007: Improving QPE and very short term QPF: An initiative for a community-wide integrated approach. *Bull. Amer. Meteor. Soc.,* **88,** 1899–1911.

Weng, F. W., L. Zhao, R. Ferraro, G. Pre, X. Li, and N. C. Grody, 2003: Advanced microwave sounding unit (AMSU) cloud and precipitation algorithm. *Radio Sci.,* **38,** 8068–8079.

Xu, L., X. Gao, S. Sorooshian, and P. A. Arkin, 1999: A microwave infrared threshold technique to improve the GOES precipitation index. *J. Appl. Meteor.,* **38,** 569–579.

Zeng, L., and G. Levy, 1995: Space and time aliasing structure in monthly mean polar-orbiting satellite data. *J. Geophys. Res.,* **100,** 5133–5142.

Zeweldi, D. A., and M. Gebremichael, 2009: Evaluation of CMORPH precipitation products at fine space–time scales. *J. Hydrometeor.,* **10,** 300–307.